Daniel Dennett on Virtual Machines

Aaron Sloman
http://www.cs.bham.ac.uk/~axs
School of Computer Science
The University of Birmingham

Related slides on several topics, including virtual machinery and consciousness, the nature of science, the fallacy of symbol grounding theory, mathematical capabilities as products of biological evolution, the nature of emotions, architectures for intelligent animals and machines, and more, are available online at
http://www.cs.bham.ac.uk/research/cogaff/talks/

Especially closely related to this presentation:

http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk86
Talk 86: Supervenience and Causation in Virtual Machinery

http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk84
Talk 84: Using virtual machinery to bridge the "explanatory gap"
Or: Helping Darwin: How to Think About Evolution of Consciousness
Or: How could evolution (or anything else) get ghosts into machines?

(Everything I do is work in progress. Comments and criticisms welcome.)

Background: The importance of virtual machinery

I first learnt about virtual machines when I visited Edinburgh University for a year (1972-3) as SRC fellow, at the invitation of Prof Bernard Meltzer, to learn about AI, while still a philosophy lecturer at Sussex University.

That year, I read Wegner (1968), which made it obvious to me that the idea of a virtual machine was very important for philosophy, including philosophy of science and philosophy of mind, as I later pointed out in several places in:

> The Computer Revolution in Philosophy: Philosophy Science and Models of Mind (Sloman, 1978)
> `http://www.cs.bham.ac.uk/research/projects/cogaff/crp/`

However, at that time I over-estimated the power of Turing machines because I had not realised their limitations, later acknowledged in: "Beyond Turing Equivalence" (Sloman, 1996):

> "There are at least three types of process that might be involved in human brains that would not necessarily map on to a single Turing machine: (a) asynchronous parallel processes with independently varying speeds, (b) continuous, or at least non-discrete processes and (c) physical processes that are not known to fit (or not to fit) the computational model, e.g. chemical processes in the brain."
> `http://www.cs.bham.ac.uk/research/projects/cogaff/0-INDEX96-99.html#1`
> (Presented at the Turing90 Colloquium, Sussex University, April 1990.)
>
> That required me to generalise the notion of virtual machinery, e.g. in Sloman (1993).

Two other philosophers who emphasise the importance of virtual machines are John Pollock (2008) and Daniel Dennett e.g. in Dennett (1991a). (They are not the only ones.)

John Pollock's characterisation is important, though he oversimplifies by not acknowledging some of the diversity and complexity of virtual machines and their relations to physical machines. However, he does not deny their existence.

Most of what Dennett writes about virtual machines is insightful, but he sometimes seems to doubt that they exist and cause things to happen, as illustrated below. Comments and corrections welcome.

Dennett on Virtual Machines (1)

Dennett writes more about virtual machines than most people.
E.g. there are several index entries for the phrase in *Consciousness Explained* (1991a).

Much of the time I think I agree with Dennett, and then he surprises me...

> In "Heterophenomenology reconsidered" Dennett (2007), he writes about: *"standard practice in computer circles, where virtual machine talk is allowed to pass for the truth with no raised eyebrows, no need for the reminder that it is, officially, just metaphorical."* (Footnote 10, p 258.)

No: when software engineers talk about what went wrong in a virtual machine, and what needs to be fixed to make it work properly, they are **not** just talking in metaphors.

> **It could be the literal truth that a decision taken by a virtual machine failed to take account of some information, or that some out of date information was mistakenly treated by the machine as true because it failed to perform some test.**

This is not misleading metaphorical shorthand, for some much more complex statement about what the transistors, or electrons did or did not do.

There's a big difference between

1. a disaster caused by some physical component going wrong, and
2. a disaster caused by information being misinterpreted, or a condition going unchecked, or use of a faulty rule, e.g. to evaluate the relative importance of two goals.

Virtual machine states and processes can exist in computers just as ignorance, poverty, crime, economic inflation, and a desire for change can exist in a social group.

See `http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk86`

Dennett on Virtual Machines (2)

In *Consciousness Explained* he writes, in connection with virtual machines:

> "In a von Neumann machine, you just 'load' the program off a disk into the main memory and the computer thereby gets an instant set of new habits...." (p. 218)

That's true if all you are doing is adding a new application virtual machine (a new software package) to an already fully installed computing system, e.g. adding a game program to an already running PC.

But the process can be very different (and far from "instant") if you acquire a computer without an operating system and want to start installing software.

Installing an operating system typically involves installing a large number of different software components that perform different functions required by the operating system.

NOTE: That was not required for the very earliest computers: individual programs could be added to them and then run. This was done in different ways at different times, e.g.

– loading the program into memory one bit at a time via switches on a panel

– reading a punched paper tape or deck of punched cards into the computer memory, via a special machine for reading such media (though software to drive the reader had to be added first).

NOTE: Even with an operating system installed, not all software will be runnable on it: e.g. you cannot simply load and run Microsoft Office on a computer running Linux as operating system.

(OpenOffice is a very good replacement for many users, but the virtual machine it provides is different.)

If you want the machine to see and manipulate things in the environment you may have a choice of cameras and arms, but the software needed to drive them in different operating systems may be different.

Dennett on Virtual Machines (3)

This quote from Dennett is better:

> "A virtual machine is What you get when you impose a particular pattern of rules (more literally: dispositions or transition regularities) on all that plasticity."
> *Consciousness Explained* p. 211

The idea of a particular set of dispositions, or transition regularities, described at a virtual machine level, is important, and is discussed more fully in other presentations here:

`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/`

But the pattern of rules need not be imposed: the rules may be grown "organically" by the information processing system interacting with its environment –

> like an animal growing new perceptual or problem-solving competences or a human learning a language. Chappell and Sloman (2007)

Moreover, Dennett sometimes suggests that the kind of virtual machine required for a human mind is a single, serial, language-using process implemented in a highly parallel collection of neurones.

That kind of "atomic state functionalism" is much too simple,

> for reasons described in Sloman (1993, 2009) (and other places):
>
> human minds, and other animal minds (like modern operating systems), need to use multi-functional virtual machines with many parallel sub-systems performing various perceptual tasks, motor control tasks, learning tasks, attention control tasks, and many more (Sloman, 1978, Ch 6).

Dennett on Virtual Machines (4)

The following is not quite correct, in general:

> "A virtual machine is a temporary set of highly structured regularities imposed on the underlying hardware by a program: a structured recipe..."
> (Densmore and Dennett, 1999)

The regularities do not have to be **temporary**: sometimes frequently used virtual machine functionality is directly supported by dedicated hardware.

A hardware floating-point unit is a physical machine, but when it is used by a running virtual machine (e.g. a virtual machine doing calculations) it provides an efficient implementation of part of the virtual machine.

However, in the majority of cases the virtual machinery that runs on a computer is temporary insofar as it can be modified, turned off, or replaced.

Dennett on Virtual Machines (5)

All of the above leaves out the important question:

Can events and processes in VMs be causes, producing effects?

Densmore and Dennett clearly seem to say yes:

> *"If you want to explain why Churchland gets the answers to the long division problems right so often, why certain problems take him longer than others, and why his pencil-pushing behavior produces the patterns of marks on the paper that it does, then the level to which you must ascend to explain is the level at which he is hand simulating the long-division machine. If instead what you want to explain are some other regularities in his behavior, such as his humming or whistling while he works, or his periodic lapses into dreamy-eyed grinning, or his muttered sub-vocalizations, then, according to Dennett, you had best descend to a somewhat lower level, but not – if you actually want to explain these patterns – all the way down to the level of the recurrent PDP networks, which are at least one level too low into the trees to permit us to see the woods we are interested in."*

Here they are talking about the reality, and the causal powers, of VM components as a software engineer would.

These components are not merely hypothesised by someone adopting the "intentional stance" on the assumption of rationality.

The existence of a particular VM running in some complex machine is a matter of fact, not a mere locutionary convenience, or metaphor.

How that fact can be established is another matter: it can be very difficult if you are not the designer of the machine.

It's also very hard for biologists studying animals: doing science can be a slow process (Lakatos, 1980).

Dennett on Virtual Machines (6)

Suppose you are shown X, then later you are sure you saw Y, Dennett (1991a) contrasts two explanations of what happened, then asserts that the distinction is bogus:

- Orwellian **O**: You saw X and experienced X, very briefly, and stored a memory of the experience; but immediately your stored memory of that experience was changed so that you remember seeing Y.
- Stalinesque **S**: Your visual subsystem got information about X, but some malfunction caused the information to be altered so it referred to Y. That was then stored, and you remember seeing Y.

Dennett and Kinsbourne (1992) claim that there is no such thing as a theatre of consciousness (ToC), only "multiple drafts" of received information (MD), so there cannot be any empirically detectable difference between **O** and **S**, and so the idea of a difference is incoherent.

The argument is criticised in Todd (2009).

From our point of view the information that you experience and are capable of reporting must reach a VM subsystem from which it can be accessed by the "meta-management" subsystem which has meta-semantic competences and also communicative competences: these are portions of the virtual machine that is your mind.

So it does make sense to distinguish two sources of error about what you experienced:

- (O) the information reached the meta-management subsystem, where it was explicitly noted, and then subsequently all records were altered, or
- (S) the information was somehow mangled in earlier stages of processing and was already incorrect by the time it reached the meta-management subsystem.

If required, we can build working models showing this distinction.

For more on meta-management see this talk (and others on the list):
`http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk87`

Dennett on Virtual Machines (7)

Note added: 29 Nov 2010 – after I wrote to Dennett

In a recent paper, Dennett (2009) writes about VMs exactly as I do: they exist, they do things, they have causal powers, and they can explain how many psychological and social/cultural phenomena are possible, including the existence of words in a shared language.

In that paper there is no suggestion that talk about VMs is “allowed to pass for the truth” though officially, just metaphorical.”

The reality of VM contents

Dennett's (occasional?) metaphysical squeamishness

- For people designing, debugging, extending, improving, managing, and using complex systems, virtual machine components, events, states, and processes are not "officially, just metaphorical" as Dennett suggests: e.g. they can cause planes to crash.

  Or to land safely in a fog.

- Many sophisticated philosophers who are familiar with mental phenomena and VMs (e.g. they use word processors, email systems, web-browsers, etc.) deny that there are events and processes that are not accessible to physical methods of detection and measurement, but can exist and explain things that are publicly observable.

  Adopting a useful stance (e.g. the intentional stance) is a poor alternative to investigating mental states and processes with properties that are introspectible and with explanatory causal powers that are not introspectible – as we should expect of virtual machine states, events and processes.

- Cautiousness about the existence of such things as beliefs, desires, intentions, preferences, inclinations, imaginings, reminiscences, twinges, and the things that have been referred to as qualia is mistaken when they are understood as occurring in VMs.

- As a result of advances over the last half century in the science and technology of computing systems, in which non-physical virtual machines play increasingly important roles, we can now look again at products of biological evolution that have similar properties.

  For more details regarding biological and non-biological virtual machinery see the references below.

The unnoticed relevance to Cognitive Science

Many cognitive scientists have heard about virtual machines (VMs) contrasted with physical machines (PMs)

Also philosophers, psychologists, neuroscientists, AI researchers, computer scientists, software engineers.

But there are differing views about how to define "virtual machine" and about the relationships between the two sorts of machine, e.g.

- Is every VM **identical** with some PM that implements or realises it?
- Can events in VMs be **causes,** or only **effect**s (only "epiphenomenal")?
- Do VMs exist or are they just **a convenient fiction**, a manner of speaking?

  (Dennett: a "centre of narrative gravity".)

If it is agreed that virtual machines and their contents can be causes (as implicitly accepted by people who think poverty can cause crime), then saying that they don't exist, or that they are just convenient fictions is pointless.

Is poverty just a convenient fiction???

Even outstanding thinkers who know about VMs sometimes forget them.

E.g. Newell and Simon's "Physical symbol system" hypothesis is wrong: they are really talking about physically implemented symbol systems in virtual machines.

(The symbols manipulated in a Lisp program are not physical objects.)

Centres of gravity and point masses (1)

ADDED: 19 Nov 2010 [NB: this is only about rigid objects.]

On (re-?)reading Dennett's "Real Patterns" (Dennett, 1991b) today, I was very surprised to notice an apparent failure to distinguish an object's centre of gravity (which exists, and is computed from its distribution of mass) from the mythical Newtonian "point masses" used in astronomical calculations. He writes (p.27)

> "It is amusing to note that my analogizing beliefs to centers of gravity has been attacked from both sides of the ontological dichotomy, by philosophers who think it is simply obvious that centers of gravity are useful fictions, and by philosophers who think it is simply obvious that centers of gravity are perfectly real:"

And then quotes Peter Smith as saying

> The trouble with these supposed parallels . . . is that they are all strictly speaking false, although they are no doubt useful simplifications for many purposes. It is false, for example, that the gravitational attraction between the Earth and the Moon involves two point masses; but it is a good enough first approximation for many calculations.

Dennett invites us to "Compare this with Fred Dretske's equally confident assertion of realism:"

> "I am a realist about centers of gravity. . . . The earth obviously exerts a gravitational attraction on all parts of the moon – not just its center of gravity. The resultant force, a vector sum, acts through a point, but this is something quite different. One should be very clear about what centers of gravity are before deciding whether to be literal about them, before deciding whether or not to be a center-of-gravity realist.

He seems to think Smith and Dretske are attacking him by saying opposite and contradictory things. Dennett does not notice that what they both say is true, and there is no contradiction.

...Continued...

Centres of gravity and point masses (2)

Dennett seems to have forgotten that Newton had proved a theorem, well known to physicists and mathematicians, but presumably not to all philosophers:

> If the inverse square law applies to tiny particles, then even a large spherically symmetrical mass also attracts masses external to its surface, even close up, exactly as if all its own mass were concentrated at its center. Thus Newton gave a mathematical justification, otherwise lacking, for applying the inverse square law to large spherical planetary masses as if they were tiny particles.
> (Quoted in `http://en.wikipedia.org/wiki/Newton's_law_of_universal_gravitation`)

The proof of this theorem (that a lengthy and and difficult calculation gives the same result as a much simpler one) does not imply that point masses actually exist, or even that they could exist (for they would have to be infinitely dense). But it enormously simplifies many calculations, where approximately spherical objects are involved.

In Newtonian physics, for objects that are perfectly spherical with a uniform distribution of mass, the theorem is not just a useful approximation, or a fiction, for it is exactly true.

That is consistent with the physical impossibility of point masses: they are indeed fictions!

In contrast, Dretske was talking not about point masses but about centres of gravity, and saying that they exist, for they are not point masses, but locations, which can be computed mathematically from the distribution of mass of any object, not just a sphere.

Existence of the centre of mass of **O** at **t** is just like existence any other location.

> Strictly this discussion should have been about “centres of mass”, not “centres of gravity”, which are different. The centre of gravity is not intrinsic to an object, but depends on the gravitational field it is in, though in many cases the centre of mass and centre of gravity will coincide or be very close.

Comments and criticisms welcome

If anyone reading this thinks I have got things wrong, please let me know – with as much evidence (e.g. references with quotations) as possible please.

Email a.sloman [AT] cs.bham.ac.uk

A revised version of this document will include acknowledgements.

In November 2010 Dennett kindly sent me a copy of Dennett (2009).

Further Reading

A debate about VMs. (The paper by Densmore and Dennett is referred to several times above).

Shannon Densmore & Daniel Dennett, "The Virtues of Virtual Machines", *Philosophy and Phenomenological Research*, 59, 3, Sep, 1999, pp. 747–761, http://www.jstor.org/stable/i345616,

P. M. Churchland, "Densmore and Dennett on Virtual Machines and Consciousness", *Philosophy and Phenomenological Research*, 59, 3, September, 1999, pp. 763–767, http://www.jstor.org/stable/2653794,

Papers and presentations on the Cognition and Affect & CoSy web sites expand on these issues, e.g.

- A. Sloman & R.L. Chrisley, (2003),
  Virtual machines and consciousness, in *Journal of Consciousness Studies*, 10, 4-5, pp. 113–172,
  http://www.cs.bham.ac.uk/research/cogaff/03.html#200302
- A. Sloman, R.L. Chrisley & M. Scheutz,
  The Architectural Basis of Affective States and Processes, in *Who Needs Emotions?: The Brain Meets the Robot,* Eds. M. Arbib & J-M. Fellous, Oxford University Press, Oxford, New York, 2005.
  http://www.cs.bham.ac.uk/research/cogaff/03.html#200305
- A. Sloman and R. L. Chrisley,
  More things than are dreamt of in your biology: Information-processing in biologically-inspired robots, *Cognitive Systems Research*, 6, 2, pp 145–174, 2005,
  http://www.cs.bham.ac.uk/research/cogaff/04.html#cogsys
- A. Sloman
  The well designed young mathematician. In *Artificial Intelligence* (2008).
  http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0807
- "What's information?" http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#905
- Presentations http://www.cs.bham.ac.uk/research/projects/cogaff/talks/

M. A. Boden, 2006, *Mind As Machine: A history of Cognitive Science* (2 Vols), Oxford University Press
(Look up "virtual machine" in the index.)

See also: http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk86
Talk 86: Supervenience and Causation in Virtual Machinery

Related Presentations

- August September 2010: Invited Talk at SAB2010, Clos-Lucè, Amboise, France, 28th Aug 2010

  Also: Talk for BPS Consciousness and Experiential Psychology Workshop Oxford 10 Sep 2010
  Helping Darwin: How to Think About Evolution of Consciousness
  How could evolution (or anything else) get ghosts into machines?
  http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk84
- Sept 2009: Virtual Machines and the Metaphysics of Science (AHRC workshop)
  http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#mos09
- July 2009: Presentation at Cognitive Science Conference, Amsterdam, 2009

  What Cognitive Scientists Need to Know about Virtual Machines
  http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#cogsci09
- 10-12 November 2008: Workshop on Philosophy and Engineering (WPE'08)
  Royal Academy of Engineering, London

  Extended abstract: Virtual Machines in Philosophy, Engineering & Biology
  http://www.cs.bham.ac.uk/research/projects/cogaff/08.html#803
  Slides here http://www.cs.bham.ac.uk/research/projects/cogaff/talks#wpe08
- 16 Oct 2008: School of Computer Science Seminar, Birmingham

  Why virtual machines really matter – for several disciplines
  http://www.cs.bham.ac.uk/events/seminars/seminar_details.html?seminar_id=560
- 21 Oct 2008: The Great Debate, Newcastle

  What can biologists, roboticists and philosophers learn from one another? (Unnoticed connections)
  http://thegreatdebate.org.uk/UnnoticedConnections.html
- 1-2 Nov 2008: Weekend course Mind as Machine, Oxford

  Why philosophers need to be robot designers
  http://oxfordphilsoc.org/

I have been writing about the importance of virtual machines intermittently since 1978 (*The Computer Revolution in Philosophy*)

Earlier talks used the slides in this (PDF) presentation
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#virt

Partial Bibliography

References

Chappell, J., & Sloman, A. (2007). Natural and artificial meta-configured altricial information-processing systems. *International Journal of Unconventional Computing*, *3*(3), 211–239. (http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609)

Dennett, D. C. (1991a). *Consciousness explained*. London and New York: Penguin Press.

Dennett, D. C. (1991b, Jan.). Real Patterns. *Journal of Philosophy*, *88*(1), 27–51.

Dennett, D. C. (2007). Heterophenomenology reconsidered. *Phenomenology and the Cognitive Sciences*, *6*(1-2), 247–270. (DOI 10.1007/s11097-006-9044-9)

Dennett, D. C. (2009, Aug). The Cultural Evolution of Words and Other Thinking Tools. In *Cold Spring Harb Symp Quant Biol published online.* Cold Spring Harbor Laboratory Press.

Dennett, D. C., & Kinsbourne, M. (1992). Time and the Observer. *Behavioral and Brain Sciences*, *15*, 183–247. (2)

Lakatos, I. (1980). Falsification and the methodology of scientific research programmes. In J. Worrall & G. Currie (Eds.), *Philosophical papers, Vol I* (pp. 8–101). Cambridge: Cambridge University Press.

Pollock, J. L. (2008). What Am I? Virtual machines and the mind/body problem. *Philosophy and Phenomenological Research.*, *76*(2), 237–309. (http://philsci-archive.pitt.edu/archive/00003341)

Sloman, A. (1978). *The computer revolution in philosophy*. Hassocks, Sussex: Harvester Press (and Humanities Press). (http://www.cs.bham.ac.uk/research/cogaff/crp)

Sloman, A. (1993). The mind as a control system. In C. Hookway & D. Peterson (Eds.), *Philosophy and the cognitive sciences* (pp. 69–110). Cambridge, UK: Cambridge University Press. (http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#18)

Sloman, A. (1996). Beyond turing equivalence. In P. Millican & A. Clark (Eds.), *Machines and thought: The legacy of alan turing (vol i)* (pp. 179–219). Oxford: The Clarendon Press. ((Presented at Turing90 Colloquium, Sussex University, April 1990. Also Cognitive Science technical report: CSRP-95-7))

Sloman, A. (2009). What Cognitive Scientists Need to Know about Virtual Machines. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1210–1215). Austin, TX: Cognitive Science Society. (http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#901)

Todd, S. J. (2009). A Difference That Makes a Difference: Passing through Dennett's Stalinesque/Orwellian Impasse,. *Br J Philos Sci*, *60*, 497–520. (http://bjps.oxfordjournals.org/cgi/reprint/60/3/497)

Wegner, P. (1968). *Programming Languages, Information Structures, and Machine Organization*. New York: McGraw Hill.